

Multimedia Systems

Lecture – 6

By

Dr. Priyambada Subudhi
Assistant Professor
IIIT Sri City

- [illegible]

How UTF -8 encoding works

- The Unicode encoding method **UTF-8** solves these problems of UTF-16 and UTF-32. It works like as follows.
 - Up to character number 128, the regular ASCII value is used (so for example A is 01000001)
 - For any character beyond 128, UTF-8 separates the code into two bytes and adding '110' to the start of first byte to show that it is a beginning byte, and '10' to the start of second byte to show that it follows the first byte.
 - So, for each character beyond number 128, you have two bytes:
[110xxxxx] [10xxxxxx] [110010 1][100001 01]
 - And you just fill in the binary for the number in between:
[11000101][10000101] (that's the number 325 -> 00101000101)

- This works for the first 2048 characters.
- For characters beyond that, one more '1' is added at the beginning of the first byte and a third byte is also used:

[1110xxxx] [10xxxxxx] [10xxxxxx]

- This gives you 16 spaces for binary code. In this manner, UTF-8 goes up to four bytes:

[11110xxx] [10xxxxxx] [10xxxxxx] [10xxxxxx]

- How the Devanagari character क, with code point 2325 can be represented in UTF-8, UTF-16 and UTF-32.

Text File Formats

- Text can be in the form of plain text or rich text.

Plain Text:

- Plain text is a text with no styles attached or embedded with it.
- It is just a plain text i.e. simple text.
- It was created by American Hardware Designers in the 60s 70s. And the files with plain text are saved with .txt or .TXT extension.

Advantages of plain text -

- if you write something in plain text then whenever you open that file on any machine then the end result will be same that is it will show the text as it is on every system. Unlike rich text if you format the text and add the styles to it then it might create problem if the end person doesn't have suitable app for it then it will distort its styles as a result it will leave a bad impression on the user.
- It is fast and flexible as it doesn't uses styles and formatting therefore it is fast and flexible.
- file size of plain text format is less as compared to rich text format.

Disadvantages of plain text -

- In the todays era almost 90% of the word processor uses rich text format.
- It looks very basic i.e. without styles.
- To read the plain text it becomes boring sometimes.

RICH TEXT :

- Rich text is a text with styles attached or embedded with it so whenever you are copying the rich text and pasting to editor which support rich text then it will automatically include all the styles formatting spacing etc.
- Rich text format was created by Microsoft in 80s and they discontinued the development of rich text in 2008. Rich text files are usually save with .rtf or .RTF extension.

Advantages of rich text -

- In almost 90% of word processors supports the rich text format.
- It looks interesting that a text with styles rather than simple text as it includes styling and formatting as a result developers learn how to style and format the text.
- To keep the end user connected to the text without getting bored.

Disadvantages of rich text -

- If the rich text file is opened in a word processor which does not support rich text then it doesn't look good to the user as the formatting might not be supported by the word processor.
- File size is large as compared to plain text as it includes styles and formatting information with it.
- It is also a bit complex to add the styles and formatting as it takes more time compared to plain text.